

Tilburg University

De voorspellende overheid

Leenes, Ronald

Published in:
Bestuurskunde

DOI:
[10.5553/Bk/092733872016025001008](https://doi.org/10.5553/Bk/092733872016025001008)

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Leenes, R. (2016). De voorspellende overheid: Transparantie is noodzakelijk, maar hoe? *Bestuurskunde*, 2016(1), 38-43. <https://doi.org/10.5553/Bk/092733872016025001008>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

De voorspellende overheid

Transparantie is noodzakelijk, maar hoe?*

Ronald Leenes

Data zijn het nieuwe goud, ook in de publieke sector. De overheid realiseert zich dat meer data de effectiviteit en efficiëntie van overheidsbeleid kunnen vergroten. Big data is goed te vergelijken met een zwarte doos, er gaat van alles in en er komt iets uit zonder dat duidelijk is wat het verband tussen in- en uitgang is. Dat is ontoelaatbaar voor voorspellende en beslissende systemen in de publieke sector. Transparantie en verantwoording zijn noodzakelijk in het licht van beginselen van de democratische rechtsstaat, de vraag is alleen hoe. Dit artikel inventariseert een aantal factoren die relevant zijn voor deze vraag.

Inleiding

Big data laat zich moeilijk definiëren (Floridi, 2012). Veelal wordt ermee bedoeld op moderne ICT-toepassingen waarin de veelheid, variëteit en vluchtigheid van gegevens centraal staan. Andere omschrijvingen benadrukken de nieuwe methoden om met data om te gaan, zoals *machine learning* en geavanceerde statistiek, en de nieuwe toepassingsmogelijkheden, zoals voorspellen, als kenmerkend voor big data. In weer andere perspectieven staat big data als mythologie centraal. De inzet van immer meer data wordt hierin gezien als middel tot een groter waarheidsgehalte, grotere objectiviteit en nauwkeurigheid (boyd & Crawford, 2012). Deze gedachte wordt treffend verwoord als ‘at sufficient scale, data is enough; “statistical algorithms find patterns where science cannot”’ (Anderson, 2008). Het is met name dit laatste perspectief dat in deze bijdrage centraal staat. De in de grote dataverzamelingen ontwaarde patronen voeden en vormen algoritmen die worden gebruikt om voorspellingen te doen en beslissingen te nemen. Deze zogenaamde *governing algorithms* verkrijgen, zoals de naam al doet vermoeden, al snel *agency* (ze beslissen, bevatten *biases*, ‘het systeem zegt ...’, et cetera) en invloed (de beslissingen worden feitelijk door het algoritme genomen en eventuele rechtsgevolgen treden in zonder menselijke tussenkomst). Bovendien zijn ze veelal ondoorgrondelijk omdat ze het gevolg zijn van statistische processen en *machine learning* en niet zozeer van gekende of bekende causale verbanden (Ziewitz, 2016). Vanuit rechtsstatelijk perspectief zijn de *agency* en intransparantie van *governing algorithms* in de publieke sector problematisch. Legitimiteit en beginselen van de democratische rechtsstaat¹ vereisen immers dat beslissingen van de overheid transparant zijn en gemotiveerd tot stand komen (Reussing,

* Prof. dr. R.E. Leenes is hoogleraar aan de Tilburg University.

1 Hieronder vallen het vrijheids-, gelijkheids-, rechtszekerheids- en het democratiebeginsel, alsmede het beginsel van de dienende overheid.

2011), hetgeen bij *governing algorithms* niet evident het geval is. De *agency* van algoritmen leidt er bovendien potentieel toe dat ze ‘macht’ verkrijgen omdat ze autonoom beslissingen nemen. Die macht zal moeten worden gecontroleerd. Transparantie en in het verlengde daarvan, publieke verantwoording, gelden als instrumenten om machtsconcentratie te beteugelen (Bovens, 2005) en kunnen de legitimiteit van het overheidshandelen middels big data mogelijk vergroten.

Big data staat relatief nog in de kinderschoenen, zeker in de publieke sector (Van der Weerd & De Vries, 2014). Het is van belang om juist nu na te denken over transparantie en publieke verantwoording om ervoor te zorgen dat het potentieel van big data op verantwoorde wijze wordt ontwikkeld. Deze korte bijdrage poogt het belang van een discussie over transparantie en publieke verantwoording rond *governing algorithms* duidelijk te maken en rijkt enige factoren aan die een rol kunnen spelen bij de invulling van dergelijke transparantie. In dat licht is enige ter-reinafbakening noodzakelijk.

Voorspellen met data

Big data toepassingen zijn te onderscheiden in fasen, van selectie van te verzamelen data, via dataverzameling, analyse en modelbouw, tot toepassing en evaluatie. Big data begint met gegevens. In de publieke sector zullen die vaak betrekking hebben op personen. Deze gegevens kunnen door het individu zelf zijn verstrekt, bijvoorbeeld via formulieren. In toenemende mate worden ze echter ook indirect verkregen door observatie (bijvoorbeeld door middel van camera's of sensoren in smartphones), door deductie uit bestaande data (bijvoorbeeld door de optelling van uitkeringen die een burger van een overheidsinstantie heeft ontvangen) en door afleiding op basis van de koppeling van verschillende gegevensverzamelingen. Zo is de vermoedelijke etnische achtergrond van migranten bijvoorbeeld af te leiden uit de vakantiebestemming en duur van de vakantie (Olsthoorn, 2014). De vier vormen waarmee gegevens over personen in het 'systeem' terechtkomen, laten een afname van de betrokkenheid (en doorgaans ook van bewustzijn) van het individu zien. Aangezien indirect verkregen gegevens buiten de burger zelf om worden verkregen, verliest deze dus ook de grip op die gegevens en wat daarmee gebeurt. Dit beïnvloedt zijn machtspositie en mogelijk ook de mogelijkheden om zijn rechten te bewaken en te effectueren.

In de wereld van big data spelen indirect verkregen gegevens een centrale rol. Interessante (nieuwe) kennis ligt, zo is de verwachting, juist besloten in de koppeling van gegevens die voor andere doelen zijn verzameld. De Belastingdienst maakt bijvoorbeeld gebruik van de kentekenfoto's die veel parkeergarages gebruiken om uitrijden te vergemakkelijken om fraude door leaserijders te detecteren.

Na de verzameling van gegevens volgt de analyse. Er worden patronen en verbanden gezocht en modellen ontwikkeld die bepaalde conclusies verbinden aan de beschikbare gegevens. Deze modellen verklaren de patronen in de data, maar maken het ook mogelijk om voorspellingen te doen over nieuwe gevallen. Daarmee is het veld van de *predictive analytics* betreden. In de online marketing wordt

Ronald Leenes

informatie over surfgedrag gebruikt om voorkeuren van mensen vast te stellen en op basis daarvan te voorspellen welke producten mogelijk interessant voor hen zijn, waarna advertenties voor juist die producten worden getoond (*preferential predictions*). Ook is het mogelijk om voorspellingen te doen over de gevolgen van bepaald gedrag (*consequential predictions*), of kunnen bepaalde toekomstige opties voor het individu worden beperkt (*pre-emptive predictions*) (Kerr & Earle, 2013). Een voorbeeld van dat laatste is dat mensen op basis van hun surfgedrag als risico worden bestempeld en daardoor bijvoorbeeld op een no-fly list worden geplaatst, wat vervolgens betekent dat hun bewegingsruimte wordt beperkt. Het is van belang te beseffen dat de voorspellingen in big data toepassingen vaak zijn gebaseerd op de projectie van patronen die zijn gevonden in grote populaties, op het individu (dat mogelijk geen deel heeft uitgemaakt van de analyse). Stel dat wordt ontdekt dat mannen tussen de 18 en 23 jaar vaker betrokken zijn bij delicten die te maken hebben met alcohol in het verkeer. Deze kennis wordt vervolgens toegepast op de 22-jarige Fons om hem te kwalificeren als een potentieel risico, daarbij (mogelijk) niet in acht nemend dat Fons geheelonthouder is en mogelijk dus geen verhoogd risico voor de verkeersveiligheid oplevert (zie bijvoorbeeld Custers, Calders, Schermer, & Zarsky, 2013).

De ontwikkelde modellen en voorspellingen kunnen worden gebruikt om besluiten in de zin van de Algemene wet bestuursrecht (Awb) te nemen. In dat geval zijn de in de Awb opgenomen waarborgen van kracht en zouden transparantie, gelijkheid, hoor- en wederhoor en dergelijke moeten zijn gegarandeerd. In met name het justitie- en veiligheidsdomein worden *pre-emptive predictions* echter ook gebruikt zonder dat er formele besluiten worden genomen, terwijl er wel consequenties zijn voor het individu.² Voorbeelden hiervan zijn de plaatsing van het individu op een *no-fly list*, aangemerkt worden als verdachte van fraude en dergelijke. In die gevallen ontbreken de Awb-waarborgen en is er sprake van de eerder geschetste spanning met beginselen van de rechtsstaat.

De voorspellingen en beslissingen vinden vanuit het perspectief van het getroffen individu doorgaans plaats in een zwarte doos. Het is bij de *pre-emptive predictions* die in dit artikel centraal staan doorgaans onduidelijk welke informatie wordt gebruikt, hoe deze wordt gewaardeerd en gewogen en hoe conclusies worden getrokken. Het hele proces is ondoorzichtig en daarmee ligt schending van het fundamentele recht op een eerlijke behandeling op de loer.

Transparantie als oplossing?

Zoals gesteld zijn transparantie en verantwoording³ middelen om overheidshandelen en machtsconcentraties te beteugelen. Dat moet mijns inziens ook het geval

2 De WRR publiceert eind maart 2016 een rapport over big data, privacy en veiligheid, waarin dergelijke toepassingen aan de orde komen.

3 Ik beperk me gezien de omvang van dit artikel tot transparantie en laat een volledige beschouwing van verantwoording achterwege. Verantwoording omvat in ieder geval een zekere mate van transparantie.

zijn voor big data toepassingen. Aandacht voor transparantie is niet alleen nodig ter bescherming van de burger, maar ook omdat de betreffende algoritmen ook voor de gebruiker (de overheid) zelf niet altijd begrijpelijk zijn, waardoor ook interne controle binnen de overheid wordt bemoeilijkt.

Transparantie is noodzakelijk om de legitimiteit van het overheidshandelen te vergroten. Uiteraard kan dat, bijvoorbeeld in het justitie- en veiligheidsdomein, op gespannen voet staan met andere beginselen, zoals effectiviteit. Als precies bekend is welke gedragingen leiden tot opname op een *no-fly list*, leidt dat ongetwijfeld tot gedragsaanpassingen van terroristen. Dat is duidelijk onwenselijk. Aantasting van de effectiviteit mag echter in een democratische rechtsstaat geen reden zijn om het vereiste van transparantie en verantwoording dan maar te laten vervallen. Belangen zoals effectiviteit kunnen wel meespelen bij de beantwoording van de vraag wat zinvolle transparantie eigenlijk moet inhouden in het geval van big data. Deze vraag kan worden onderverdeeld in vier subvragen: *wie* moet transparantie bieden of verantwoording afleggen, *aan wie*, *waarover* en *waarom* (Bovens, 2005). Ik laat de vraag wie transparantie moet geven liggen en ga ervan uit dat dit het bestuursorgaan of de overheidsinstantie is die de *governing algorithms* inzet. De andere vragen zijn lastiger te beantwoorden en zijn met elkaar verweven en afhankelijk van het toepassingsdomein en de specifieke toepassing. Ik geef hieronder een paar overwegingen.

Over de waarom-vraag is hierboven in algemene zin al iets gezegd. Bovens (2005) plaatst het doel van publieke verantwoording in de context van drie (concurrerende) theoretische perspectieven: democratische controle, behoorlijk bestuur en leervermogen. Iets losser geïnterpreteerd vertegenwoordigen ze drie belangen die goed bruikbaar zijn in de context van *pre-emptive predictions*. Vanuit het perspectief van democratische controle kan worden vereist dat de overheid de behoorlijkheid en doelmatigheid van de gehanteerde technieken duidelijk maakt. Dat vereist niet per definitie een gedetailleerd inzicht in data, modellen en uitkomsten, maar wel transparantie en publieke verantwoording over het waarom van bepaalde toepassingen, de effectiviteit en de nadelige effecten. Dit maakt het mogelijk voor de toezichthouder of het parlement om in te grijpen wanneer rechten van burgers te veel op het spel staan. Ook leervermogen speelt hier een rol. Transparantie over wat wel en niet aanvaardbaar geacht is in het verleden, voorkomt mogelijk de ontwikkeling van onacceptabele toepassingen in de toekomst. Het belang van behoorlijk bestuur kan worden ingevuld als transparantie (en verantwoording) naar het getroffen individu over de behoorlijkheid van de beslissing. De drie belangen hebben (gedeeltelijk) verschillende publieken. Democratische controle richt zich op de kiezer, het parlement en de politieke bestuurder, leervermogen richt zich meer op uitvoerende instanties en bestuurders, behoorlijk bestuur richt zich primair op de burger.

In het licht van beginselen van de democratische rechtsstaat zou het uitgangspunt moeten zijn dat daar waar mogelijk het getroffen individu informatie wordt verstrekt over de gehanteerde data, modellen, en concrete toepassing, analoog

aan het Awb/Wbp-regime,⁴ zodat deze zich kan verweren en *due process* kan worden gewaarborgd. Daar waar overwegende redenen bestaan om hier van af te wijken, bijvoorbeeld omdat dit de effectiviteit van de toepassing drastisch zou verminderen of de (publieke) veiligheid in gevaar komt, is (gedetailleerde) informering van betrokkenen minder vanzelfsprekend. Kennis over de gehanteerde data en modellen maakt het immers tot op zekere hoogte mogelijk de uitkomsten te manipuleren en het doel van het systeem te ontduiken. Dit probleem moet echter niet worden overschat aangezien burgers zich in veel gevallen feitelijk niet aan gegevensverzameling kunnen onttrekken. Kentekengegevens gebruikt door parkeergarages om uitrijden te vergemakkelijken worden, bijvoorbeeld, aan de Belastingdienst verstrekt, die deze gebruikt ter detectie van fraude door leaserijders. Daar is geen ontkomen aan. *Gaming* is op dit vlak nauwelijks mogelijk. Daarnaast kan transparantie ook bijdragen aan de doelen van het systeem omdat het normconformerend gedrag van burgers kan bevorderen wanneer ze weten dat handhavend wordt opgetreden. Wanneer ontduiking als gevolg van transparantie reëel is, dan zal de verantwoording anders moeten worden ingericht. Toezicht door een onafhankelijke autoriteit (audits) is dan wenselijk. Hiermee kan tegemoet worden gekomen aan de behoefte aan controle, maar kan tevens het leervermogen van de betreffende organisatie worden vergroot doordat met andere ogen naar de invoer, processen en uitvoer kan worden gekeken.

Conclusie

Veel big data toepassingen zijn ondoorgrondelijke zwarte dozen; er gaat van alles in en wat er uit komt 'werkt', vraag niet hoe. Vaak is daar niets mis mee, de gevolgen zijn veelal beperkt. In de publieke sector moet de lat hoger liggen; de overheid heeft wat uit te leggen wanneer ze voorspellingen en beslissingen over burgers baseert op big data. Welke vorm dit moet krijgen, vereist nog veel onderzoek (Kerr & Earle, 2013).

Literatuur

- Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired* 23 June. Retrieved from http://www.wired.com/science/discoveries/magazine/16-07/pb_theory
- Bovens, M. (2005). Publieke verantwoording: Een analysekader. In W. Bakker & K. Yesilkagit (red.), *Publieke verantwoording* (pp. 25-55). Amsterdam: Boom.
- boyd, d., & Crawford, K. (2012). Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon, *Information, Communications, & Society*, 15, 662.

4 De Wet bescherming persoonsgegevens biedt een recht op inzage in de gehanteerde gegevens en de logica die ten grondslag ligt aan beslissingen (art. 35 lid 4 Wbp) en het recht niet onderworpen te worden aan geautomatiseerde beslissingen zonder menselijke tussenkomst (art. 42 Wbp).

- Custers, B., Calders, T., Schermer, B., & Zarsky, T. (eds.). (2013). *Discrimination and privacy in the information society – Data mining and profiling in large databases*. Heidelberg etc.: Springer.
- Floridi, L. (2012). Big Data and Their Epistemological Challenge. *Philosophy and Technology*, 25, 435-437.
- Kerr, I., & Earle, J. (2013). Prediction, Preemption, Presumption: How Big Data Threatens Big Picture Privacy, *Stanford Law Review Online*, 66, 65-72.
- Olsthoorn, P. (2014). *Verdacht door data. Van tandenborstels tellen naar datafuiken*. Verkregen van <http://ibestuur.nl/magazine/verdacht-door-data>
- Reussing, R. (2011). De legitimiteit van de democratische rechtsstaat: inhoudelijk of procedureel? *Bestuurswetenschappen*, 5, 52-74.
- Weerdt, C. van der, & Vries, A. de. (2014). *Dienstverlening verbeteren met Big Data: een verkenning voor gemeenten*. TNO. Verkregen via <http://kennisopenbaarbestuur.nl/rapporten-publicaties/dienstverlening-verbeteren-met-big-data/>
- Ziewitz, M. (2016). Governing Algorithms: Myth, Mess, and Methods, *Science, Technology, & Human Values*, 41(1), 3-16.